Indel detection from Whole Genome Sequencing data and association with lipid metabolism in pigs

Daniel Crespo-Piazuelo [1,2]*, Lourdes Criado-Mesas[1], Manuel Revilla [1,2], Anna Castelló[1,2], Ana I. Fernández[3], Josep M. Folch[1,2], Maria Ballester[4]

1 Plant and Animal Genomics, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB Consortium, Bellaterra, Spain, **2** Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain, **3** Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain, **4** Departament de Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Caldes de Montbui, Spain

* daniel.crespo@cragenomica.es

Abstract

The selection in commercial swine breeds for meat-production efficiency has been increasing among the past decades, reducing the intramuscular fat content, which has changed the sensorial and technological properties of pork. Through processes of natural adaptation and selective breeding, the accumulation of mutations has driven the genetic divergence between pig breeds. The most common and well-studied mutations are single-nucleotide polymorphisms (SNPs). However, insertions and deletions (indels) usually represents a fifth part of the detected mutations and should also be considered for animal breeding. In the present study, three different programs (*Dindel*, *SAMtools mpileup*, and *GATK*) were used to detect indels from Whole Genome Sequencing data of Iberian boars and Landrace sows. A total of 1,928,746 indels were found in common with the three programs. The *VEP* tool predicted that 1,289 indels may have a high impact on protein sequence and function. Ten indels inside genes related with lipid metabolism were genotyped in pigs from three different backcrosses with Iberian origin, obtaining different allelic frequencies on each backcross. Genome-Wide Association Studies performed in the *Longissimus dorsi* muscle found an association between an indel located in the C1q and TNF related 12 (*C1QTNF12*) gene and the amount of eicosadienoic acid (C20:2(n-6)).

Introduction

Pork is one of the world's most produced meat. Selective breeding in pigs has been developed in parallel to the increase and intensification of this productive sector. Over the last decades, genetic selection has notably improved meat-production efficiency in commercial pig breeds. However, this artificial selection had the unwanted drawback of reducing the pork sensorial and technological properties of meat. These modifications were driven by the reduction of intramuscular fat (IMF) content and fatty acid (FA) composition changes [1].

Commercial breeds as Landrace possess an efficient meat production with a rapid growth and a leaner carcass, but the resulting meat has lower IMF and higher polyunsaturated FAs (PUFA) content compared with some indigenous pig breeds, such as the Iberian pig [2]. The Iberian breed is characterized by its higher IMF content with a great proportion of monounsaturated FAs (MUFA) [3]. In addition, MUFA have a more oxidative stability than PUFA, improving the organoleptic properties of meat [4]. In contrast, PUFA consumption, in particular omega-3, has the beneficial role of decreasing the total cholesterol concentration, while saturated FAs (SFA) increase the risk of suffering cardiovascular diseases [5,6].

Fatty acid composition in muscle is determined by physiological conditions such as fed and fasted states [7], environmental factors such as nutrition [4,8] and genetic factors; carcass and FA composition traits in pigs that range from moderate to high heritability values [9–12].

The genetic divergence between breeds has been driven by the accumulation of mutations through processes of natural adaptation to the environment and selective breeding. Genetic mutations can be produced by base pair substitution, but also by insertion, inversion, fusion, duplication or deletion of DNA sequences. The development of next generation sequencing (NGS) technologies has improved the detection of these genomic variants. Hitherto, the most well-known variants studied with this method have been the substitutions of single nucleotide polymorphisms (SNPs), which represent almost the 80% over all the detected variants [13–15]. In contrast, insertions and deletions (indels) have been less characterized although the genome-wide ratio of indels to SNPs has been estimated as 1 indel for every 5.3 SNPs [16]. Studies in *Drosophila melanogaster* and *Caenorhabditis elegans* have determined that indels represent between 16% and 25% of all genetic polymorphisms in these species [17,18]. In addition, studies performed in humans and chimpanzees evidenced that indels instead of SNPs were the major source of evolutionary change [19–21].

As it has been described over the last decades, the most frequently found indel was the 1 base pair (bp) long [22,23]. Furthermore, a major proportion of deletions than insertions was observed in the genome of 18 mammals, with the exception of the opossum [24]. A mechanism that favours the occurrence of deletions was proposed by de Jong & Rydén [25], in which the loops formed by slipped mispairing after DNA strand breakage are trimmed off. In pigs, recent studies using whole genome sequencing (WGS) have detected the 1 bp long indel as the most frequent indel, but the deletion/insertion ratios differ [26–29].

Indels can produce frameshifts in the reading frame of a gene or modify the total number of amino acids in a protein, but they can also affect gene expression levels. In pigs, indels were found to affect backfat thickness [30] and fat deposition [31] through the alteration of gene expression, underlining the importance of these variants for animal production.

The objectives of this study were to identify indels from WGS data of Iberian and Landrace pigs, which were founders of an experimental cross (IBMAP) with productive records for FA composition, and to study the association between a selection of indels and meat quality traits in three different genetic backgrounds.

Material and methods

Ethics statement

The present study was performed in accordance with the regulations of the Spanish Policy for Animal Protection RD1201/05, which meets the European Union Directive 86/609 about the protection of animals used in experimentation. All experimental procedures followed national and institutional guidelines for the Good Experimental Practices and were approved by the IRTA (*Institut de Recerca i Tecnologia Agroalimentàries*) Ethics Committee.

Animal material and phenotypic records

The pigs used in this study belonged to the Iberian and Landrace breeds. The Iberian line, called Guadyerbas, is a unique black hairless line that has been genetically isolated in Spain since 1945 [32]. The Landrace line belonged to the experimental farm Nova Genètica S.A. (Lleida, Spain). WGS data of seven founders of the IBMAP experimental population [32], two Iberian boars and five Landrace sows, were used for indel detection. Analysis of indel segregation and association with meat quality traits were performed in 441 individuals of different backcrosses: 160 BC1_LD ((Iberian x Landrace) x Landrace), 143 BC1_DU ((Iberian x Duroc) x Duroc) and 138 BC1_PI ((Iberian x Pietrain) x Pietrain). All animals were reared in the experimental farm of Nova Genètica S.A. (Lleida, Spain). Population structure of these three backcrosses is depicted in S1 Fig.

Animals were fed *ad libitum* with a cereal-based commercial diet and slaughtered at an average age of 179.8 ± 2.6 days with an average carcass weight of 72.2 kg. Blood from founder animals was collected in 4 ml EDTA vacutainer tubes and stored at -20°C until analysis. Samples of diaphragm tissue were collected from backcrossed animals, snap-frozen in liquid nitrogen and stored at -80°C until analysis. Genomic DNA was extracted from all samples by the phenol-chloroform method [33].

At the slaughterhouse, 200 g of *Longissimus dorsi* muscle samples were collected from the three backcrosses. The IMF composition was measured with a protocol based on gas chromatography of methyl esters as described in Pérez-Enciso *et al.* [32]. In total, 20 traits were

Table 1. Descriptive statistics including mean and SD of fatty acid composition and FA indices in the *Longissimus dorsi* muscle of the merged dataset and the three backcrosses.

Group	Trait	Name	Merged dataset		BC1_DU		BC1_LD		BC1_PI	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
SFA	C14:0	Myristic acid	1.17	0.24	1.27	0.23	1.18	0.15	1.06	0.27
	C16:0	Palmitic acid	22.99	1.51	23.94	1.65	22.59	1.20	22.42	1.14
	C17:0	Margaric acid	0.27	0.10	0.25	0.10	0.27	0.07	0.30	0.13
	C18:0	Stearic acid	14.21	1.44	14.33	1.75	14.18	1.03	14.13	1.46
	C20:0	Arachidic acid	0.26	0.11	0.23	0.08	0.26	0.12	0.31	0.12
MUFA	C16:1(n-9)	7-Hexadecenoic acid	0.36	0.11	0.31	0.12	0.39	0.09	0.39	0.10
	C16:1(n-7)	Palmitoleic acid	2.61	0.50	2.81	0.52	2.50	0.39	2.53	0.52
	C17:1	Heptadecenoic acid	0.23	0.10	0.19	0.09	0.27	0.11	0.23	0.10
	C18:1(n-9)	Oleic acid	37.08	5.78	35.97	5.63	40.07	2.77	35.15	6.92
	C18:1(n-7)	Vaccenic acid	3.91	0.33	3.83	0.30	3.88	0.36	4.03	0.30
	C20:1(n-9)	Gondoic acid	0.82	0.20	0.73	0.16	0.85	0.11	0.88	0.26
PUFA	C18:2(n-6)	Linoleic acid	11.92	5.01	12.11	5.83	10.36	2.38	13.32	5.64
	C18:3(n-3)	α-Linolenic acid	0.50	0.23	0.40	0.13	0.65	0.29	0.44	0.14
	C20:2(n-6)	Eicosadienoic acid	0.51	0.14	0.43	0.12	0.54	0.12	0.57	0.14
	C20:3(n-3)	Eicosatrienoic acid	0.22	0.13	0.18	0.10	0.20	0.15	0.28	0.13
	C20:3(n-6)	Dihomo-γ-linolenic acid	0.43	0.28	0.45	0.29	0.28	0.13	0.58	0.29
	C20:4(n-6)	Arachidonic acid	2.53	2.04	2.57	1.94	1.54	0.74	3.49	2.53
Metabolic Ratios	SFA	Saturated fatty acids	38.90	2.44	40.01	3.13	38.47	1.64	38.17	1.84
	MUFA	Monounsaturated fatty acids	44.78	6.25	43.54	6.05	47.95	3.07	42.81	7.54
	PUFA	Polyunsaturated fatty acids	15.88	7.26	15.96	8.12	13.37	3.30	18.37	8.37

analysed: 17 intramuscular FAs and 3 FA metabolism indices (Table 1). Data values were normalized applying a $\log_2$ transformation when needed.

Whole genome sequencing

The whole genome of seven founders of the IBMAP population was sequenced at CNAG (National Centre for Genome Analysis, Barcelona, Spain) on an Illumina HiSeq2000 instrument (Illumina, San Diego, CA, USA). Paired-end sequencing libraries, with approximately 300 bp insert size, were generated using TruSeq DNA Sample Prep Kit (Illumina, San Diego, CA, USA). For each sample, around 40 million 100 bp-long paired-end reads were produced with an average sequencing depth of 11.7x. Whole genome sequencing files of the seven BC1_LD founders are described in Revilla *et al*. [34] and were deposited in the NCBI Sequence Read Archive (SRA) under accession nos. SRR5229970, SRR5229971, SRR5229972, SRR5229973, SRR5229974, SRR5229975 and SRR5229976.

Sequences were trimmed based on their quality using the FastQC [35] software. Then, reads were mapped against the reference genome sequence assembly (*Sscrofa10.2*) using the Burrows-Wheeler Alignment (BWA) tool [36]. Duplicated reads or those which were under a Phred-based quality score of 20 were removed. Finally, alignment result files (in bam format) were prepared for indel detection.

Indel detection and effects prediction

Several programs allow performing indel calling from WGS bam files. Following the article of Neuman *et al*. [37] on the comparison of short indel detection programs, we applied the recommended pipelines on the use of these three programs: *Dindel* (version 1.01) [38], *SAMtools mpileup* (version 0.1.19) [39], and *Genome Analysis Toolkit* (*GATK*) (version 3.4–46) [40].

The *Variant Effect Predictor* (*VEP*) (version 82) [41] tool of Ensembl (http://www.ensembl.org/) was used to quickly and accurately predict the effects and consequences of indels previously found on Ensembl-annotated transcripts [41]. Furthermore, to predict the possible effect of an indel in the secondary structure of a protein, *JPred4* [42] was used.

Finally, ten indels were selected for indel validation and association analysis if they followed any of these two criteria:

1. those start or stop variants related with lipid metabolism

2. those indels with high or moderate severity that were found at extreme frequencies in the founder animals (IB = 1 & LD≤0.2 or IB = 0 & LD≥0.8). Among this subset of 127 indels, those involved in lipid metabolism were prioritized.

Genotyping

For indel validation and association analysis, ten indels were genotyped in three experimental backcrosses: BC1_DU (n = 143), BC1_LD (n = 160) and BC1_PI (n = 138), using Taqman OpenArray genotyping plates custom designed in a QuantStudio 12K flex Real-Time PCR System (ThermoFisher Scientific, Waltham, MA, USA).

The same animals of BC1_LD and BC1_PI were genotyped with the Porcine SNP60K BeadChip (Illumina, San Diego, CA, USA), while BC1_DU samples genotypes were obtained with the Axiom Porcine Genotyping Array (Axiom_PigHDv1; Affymetrix, Santa Clara, CA, USA). Only those variants shared by both genotyping platforms were kept. A total of 38,424 SNPs remained after removing SNPs with a minor allele frequency (MAF) < 5% and SNPs with missing genotype > 5% data using *PLINK* (1.90b5 version) [43].

Genome-Wide association analysis

Genome-Wide Association Studies (GWAS) were performed between the measured phenotypes of IMF composition and the previously genotyped variants of the three backcrosses (38,424 SNPs and nine indels) along the pig reference genome assembly (*Sscrofa11.1*). The studies were conducted with *GEMMA* [44] following the mixed linear model:

$$y_{ijklm} = Sex_i + Batch_j + Backcross_k + \beta c_l + u_l + \delta_l a_m + e_{ijklm},$$

where y_{ijklm} indicates the value of the phenotypic observation in the lth individual; sex (two categories), batch (fourteen categories) and backcross (three categories) are fixed effects; β is a covariate coefficient with c being carcass weight; u_l is the infinitesimal genetic random effect and distributed as $N(0, K\sigma_u)$, where K is the numerator of the kinship matrix; δ_l represents the allelic effect, calculated as a regression coefficient on the lth individual genotype for the mth SNP or indel (values -1, 0, +1); a_m represents the additive effect associated with the mth SNP or indel; and e_{ijklm} is the random residual term. Genomic kinship was obtained selecting the "-gk 1" option in *GEMMA* software [44], which calculates a centred relatedness matrix using the genotypic information of the individuals.

GWAS were also performed individually for each one of the three backcrosses following the previously described model, except for the fixed effect of the backcross which was removed from the model.

The multiple test correction was conducted with the *p.adjust* function incorporated in R (www.r-project.org) using the false discovery rate (FDR) method developed by Benjamini and Hochberg [45]. In order to consider a SNP or an indel as significant or suggestive a cut-off was set at FDR≤0.05 or FDR≤0.1, respectively.

Results and discussion

Genome-wide detection of indels in Iberian and Landrace animals

Whole genome sequencing data of seven founders of the IBMAP population (two Iberian boars and five Landrace sows) were used for indel detection with *Dindel*, *SAMtools mpileup* and *GATK* software. *Dindel* was the program that detected the highest number of indels (3,380,221) as opposed to *SAMtools mpileup* and *GATK* (2,749,596 and 2,957,377, respectively). To reduce the rate of false positives, only indels (1,928,746) that were found in common between the three programs were considered for further analyses (Fig 1). In addition, 50,528 indels were discarded for not displaying the same genotype in at least two programs.

Repetitive elements, such as microsatellites, are short insertions or deletions that can interfere with the detection and annotation of indels. Thus, to reduce the interference of repetitive elements in the next steps, 105,783 variants were discarded if they were triallelic or the alternative allele was different among individuals for the same chromosomal position. Moreover, 141,391 indels were trimmed because they were homozygous for the alternative allele in all samples and may not be segregating in our population. Hence, we only considered the final list comprising 1,631,044 indels for further analysis (S1 Table).

In a preliminary study of our group, in which SNP calling was performed from WGS of these seven IBMAP founders, the number of SNPs identified after the quality filter was 4.9 million in the Iberian boars and 6 million in the Landrace sows. Therefore, the number of indels detected (1.6 million indels) was within the expected range (16–25%) of the total number of variants detected [13–18]. Nevertheless, another study in pigs reported that indels were less frequent than SNPs in a proportion of 1 to 10 [26].
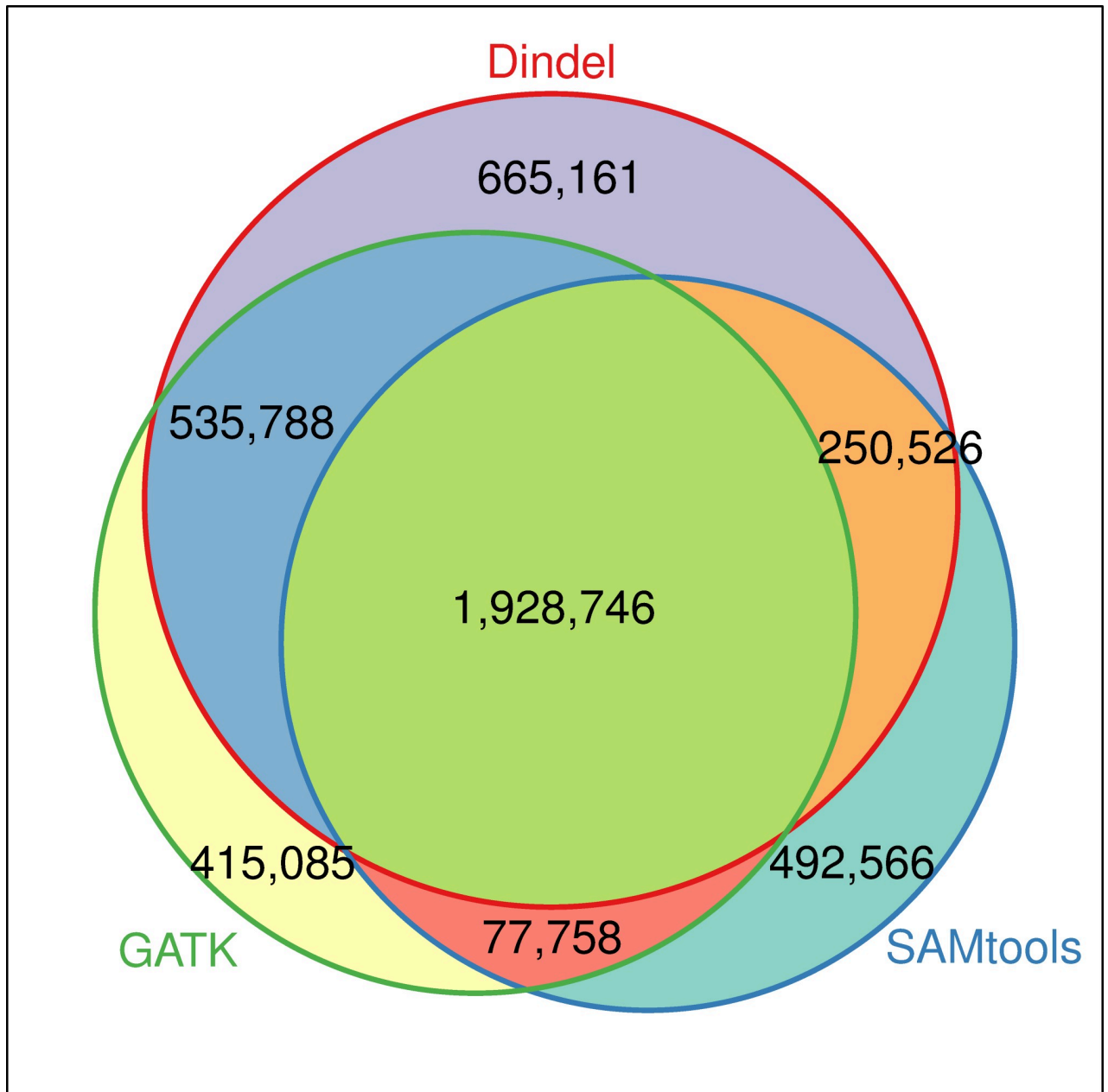
Fig 1. Weighted Venn diagram showing the number of indels shared between the three indel detection programs: *Dindel*, *Pindel* and *SAMtools* mpileup. A total of 1,928,746 indels were found in common.

The distribution of the indels found along all the *Sus scrofa* chromosomes (SSC) showed that sexual chromosomes (SSCX and SSCY) had lower density of indels than autosomes (Fig 2). Disregarding the pseudoautosomic regions, this low density of indels in the sexual chromosomes is probably caused by the low recombination rate, only possible for the X chromosome in females, and by the appearance of hemizygous recessive lethal mutations in males. In addition, males present one copy of each heterosome, and accordingly, the density of mutations in autosomes, which have two copies of each chromosome, is higher than in heterosomes. The autosome that had the highest density of indels was SSC10, while SSC1 had the lowest (Fig 2).
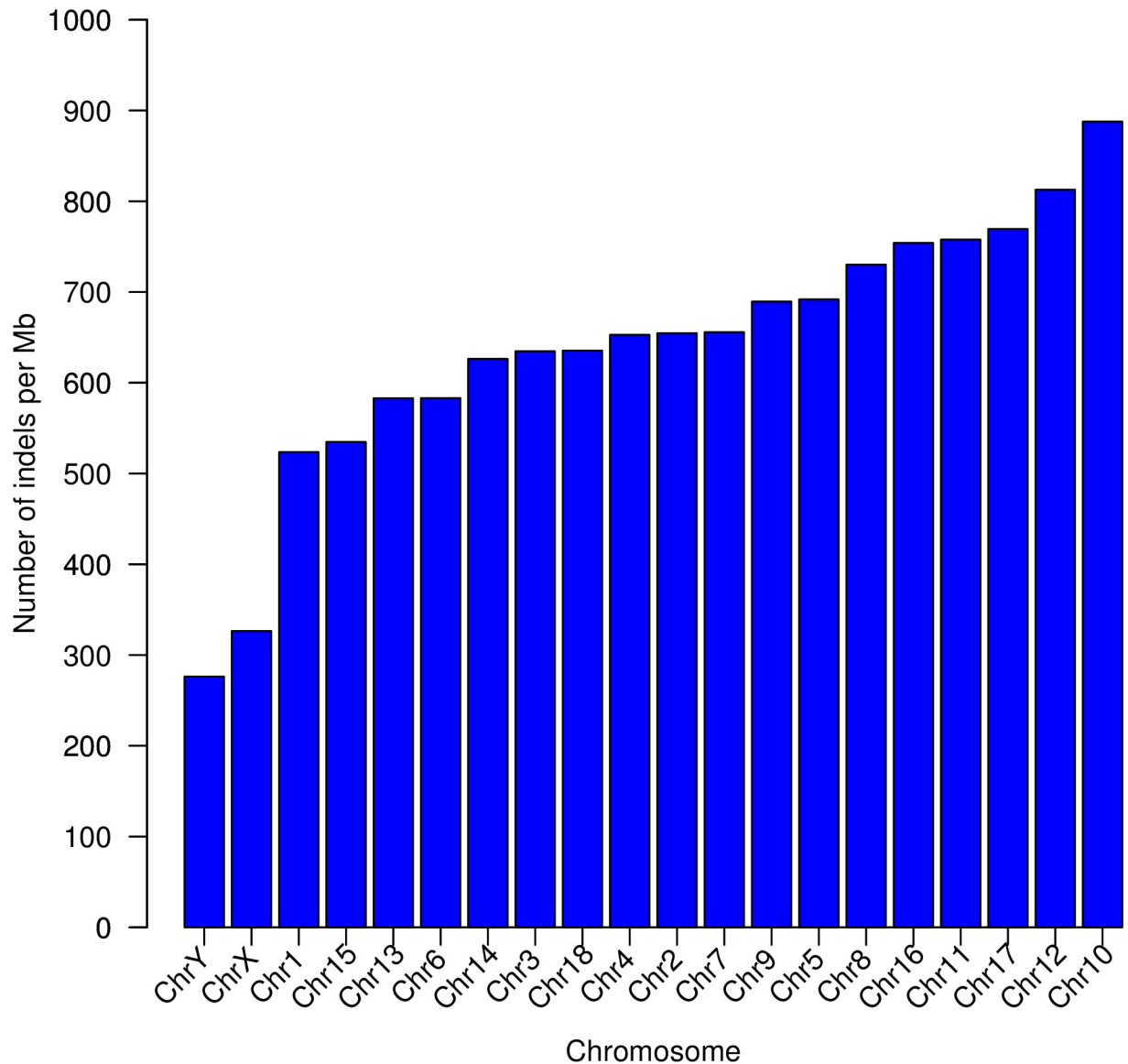
Fig 2. Distribution of the density of indels across chromosomes calculated as number of indels per Mb. Chromosomes are sorted in increasing order of density value.

In accordance with the literature, indel frequencies decreased as their length increased [27,46] and thus, 1 bp long indel was the most frequent indel found (Fig 3), either insertion or deletion [22,23]. Insertions were more frequent than deletions in single bp indels, but from the 1.6 million indels, 52.9% were deletions from 1 to 54 bp and the rest were insertions (47.1%) from 1 to 32 bp. Therefore, deletions were found to be more frequent than insertions, which has been previously reported by some other studies made in pigs [26,28] and follows the mutational mechanisms described by de Jong & Rydén (1981).

Consequence and severity predictions of the indels detected

The effects (consequence type and severity) of the 1.6 million indels were estimated by the *VEP* platform and are summarized in Table 2. Since a variant may co-locate with more than
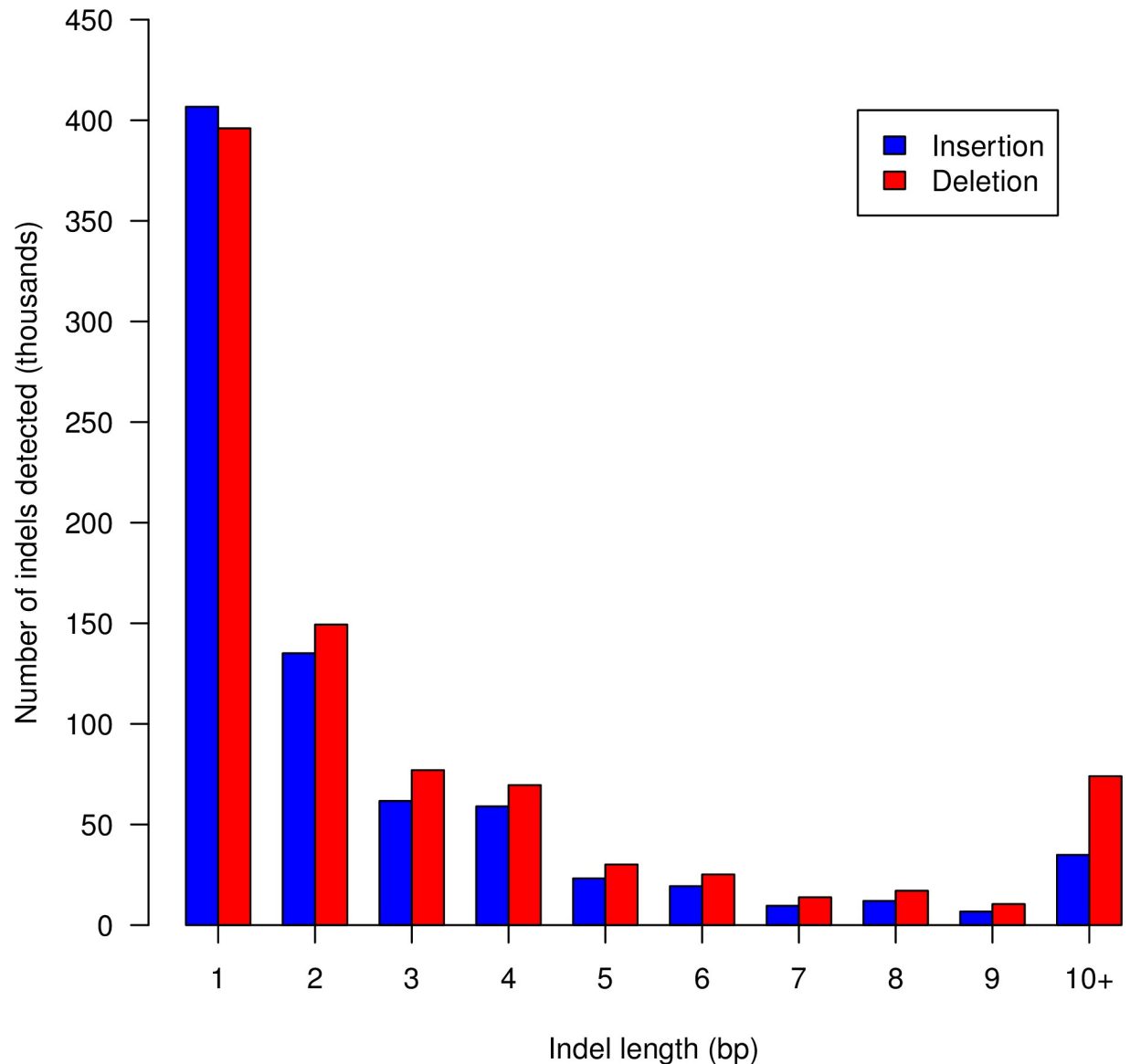
Fig 3. From the total of 1,631,044 indels detected, it is represented the quantity of them according to their length in bp. Insertions are in red and deletions are in blue.

one transcript, one line of output was provided for each instance of co-location and thus, there were more lines written (1,790,722) than indels entered (1,631,044). In addition, the total number of predicted effects was 1,809,798 as some indels can result in more than one effect in the same transcript (e.g., an indel could cause a frameshift along with a stop gained). Around the third part of the 1.6 million indels (33.1%) did not fall within intergenic regions (539,920 indels) and only 1,758 indels were inside a coding region (0.11%). Finally, the *VEP* platform classified the 1.6 million indels by their possible severity as high (1,289), moderate (561) or low (1,018) impact, and the rest of indels were considered as modifiers.

Indel selection for genotyping

From the total of indels with high and moderate impact (1,850), ten indels were selected to be genotyped in three different genetic backgrounds. These indels were chosen regarding their

Table 2. Consequences predicted by the *VEP* platform.

Consequence type	Quantity	VEP severity
Total of indels processed (input)	**1,631,044**	-
intergenic variant	1,091,124	Modifier
intron variant	506,323	Modifier
downstream gene variant	91,517	Modifier
upstream gene variant	90,313	Modifier
non coding transcript variant	11,535	Modifier
3' UTR variant	7,618	Modifier
NMD transcript variant	5,911	Modifier
splice region variant	1,443	Low
frameshift variant	1,246	High
5' UTR variant	1,112	Modifier
non coding transcript exon variant	650	Modifier
inframe deletion	359	Moderate
inframe insertion	285	Moderate
coding sequence variant	115	Modifier
splice acceptor variant	94	High
splice donor variant	74	High
mature miRNA variant	41	Modifier
start lost	14	High
stop gained	9	High
protein altering variant	6	Moderate
stop retained variant	5	Low
stop lost	3	High
incomplete terminal codon variant	1	Low

https://doi.org/10.1371/journal.pone.0218862.t002

possible consequence, if they were inside genes that could be related with lipid metabolism and/or considering their frequencies in the founder animals.

Table 3 summarizes the list of genes with indels selected for genotyping:

1. The aspartate beta-hydroxylase (*ASPH*) gene (ENSSSCG00000025087), located on SSC4, contained a predicted frameshift variant (rs691136075) with a high impact. The expression of this gene was found to be negatively correlated with insulin-stimulated sprouting in mice adipose tissue [47].

2. The calpain 9 (*CAPN9*) gene (ENSSSCG00000010182) is located on SSC14 and contained a predicted inframe deletion (rs704351652). *CAPN9* is a member of the calpain family and some of its members have been associated with body fat content and insulin resistance in human and mice [48,49]. This variant was found at extreme frequencies in the founder animals being the alternative allele (*CAPN9:c.2013_2015delGAA*) fixed in the Iberian boars.

3. The C-C motif chemokine receptor 7 (*CCR7*) gene (ENSSSCG00000017466) is located on SSC12 and contained a predicted frameshift variant (rs789030032). *CCR7* codifies for a chemokine receptor that plays a crucial role in inducing adipose tissue inflammation, insulin resistance and obesity [50,51]. The allele frequency for this indel (*CCR7:c.1142dupA*) in the Landrace sows was 0.5 while the two Iberian boars were homozygous for the reference allele.

4. The C-reactive protein (*CRP*) gene (ENSSSCG00000021186), located on SSC4, contained a frameshift variant (*CRP:c.515delT*). High levels of *CRP* has been related with overweight

Table 3. Selection of the ten genotyped indels with the alternative allele frequency in the Iberian (Freq. IB) and Landrace (Freq. LD) founders and their consequence predicted by the *VEP* platform.

Gene Name	Ensembl Gene ID	Chr.	Position[a]	Reference Allele	Alternative Allele	Freq.[b] IB	Freq.[b] LD	Consequence	Severity
ASPH	ENSSSCG00000025087[c]	4	78,502,739 (72,104,018)	T	TAGAC	0	0.1	Frameshift variant Stop retained variant	High
PEX19	ENSSSCG00000023091	4	98,087,517 (90,197,086)	G	GCAAGT	1	0.2	Frameshift variant	High
CRP	ENSSSCG00000021186	4	98,755,304 (90,783,699)	GT	G	1	0	Frameshift variant	High
C1QTNF12	ENSSSCG00000003333	6	57,988,405 (63,549,854)	ACCG	A	1	0	Inframe deletion	Moderate
CCR7	ENSSSCG00000017466	12	22,151,183 (21,868,256)	T	TA	0	0.5	Frameshift variant Stop retained variant	High
LYST	ENSSSCG00000010151	14	59,597,172 (55,547,902)	G	GACC	0	0.8	Inframe insertion	Moderate
CAPN9	ENSSSCG00000010182	14	64,251,137 (59,585,649)	GTTC	G	1	0	Inframe deletion	Moderate
JMJD1C	ENSSSCG00000010226	14	71,899,504 (66,662,629)	TCTG	T	1	0	Inframe deletion	Moderate
GZMA	ENSSSCG00000016903	16	36,388,745 (34,285,211)	G	GTGT	0	0.8	Inframe insertion	Moderate
SAMD4B	ENSSSCG00000016927[c]	16	40,171,247 (37,404,073)	TCA	T	0.25	0.7	Frameshift variant Stop gained	High

[a]Position is referred to the indel's previous nucleotide in the *Sscrofa10.2* assembly. Position in the *Sscrofa11.1 assembly* is enclosed in parentheses.

[b]Freq. stands for the frequency of the alternative allele.

[c]Novel pig genes that were orthologous with their human, mouse and cow counterpart genes.

https://doi.org/10.1371/journal.pone.0218862.t003

and obesity in human adults [52]. This variant was found fixed in the Iberian boars for the alternative allele (*CRP:c.515delT*) and the alleles of the Landrace sows were as the reference.

5. The C1q and TNF related 12 (*C1QTNF12*) gene (ENSSSCG00000003333) is located on SSC6 and contained an inframe deletion (*C1QTNF12:c.557_559delCCG*). This gene is also known as *CTRP12* and *FAM132A*. C1QTNF12 functions as an adipokine that is involved in glucose metabolism and obesity in mice [53,54]. This deletion was found at extreme frequencies in the founders being the alternative allele (*C1QTNF12:c.557_559delCCG*) fixed in the Iberian boars.

6. The granzyme A (*GZMA*) gene (ENSSSCG00000016903), located on SSC16, contained an inframe insertion (rs792025734). This gene was differentially expressed in the mesenteric adipose tissue of beef cattle with distinct gain [55]. The insertion (*GZMA: c.129_131dupGTT*) was found with a frequency of 0.8 in the Landrace sows while the Iberian boars were homozygous for the reference allele.

7. The jumonji domain containing 1C (*JMJD1C*) gene (ENSSSCG00000010226) is located on SSC14 and contained an inframe deletion (*JMJD1C:c.5964_5966delCAG*). *JMJD1C* was found in a human GWAS as a candidate gene for very low-density lipoprotein particles [56]. This variation was found at extreme frequencies in the founders being the alternative allele (*JMJD1C:c.5964_5966delCAG*) fixed in the Iberian boars.

8. The lysosomal trafficking regulator (*LYST*) gene (ENSSSCG00000010151), located on SSC14, contained an inframe insertion (rs713515754). This gene has been related with hypertriglyceridemia and anomalous lipid and FA composition in the erythrocyte

membranes of Chédiak-Higashi human patients [57]. This variation (*LYST: c.6287_6289dupCCA*) was found with a frequency of 0.8 in the Landrace sows while the Iberian boars were homozygous for the reference allele.

9. The peroxisomal biogenesis factor 19 (*PEX19*) gene (ENSSSCG00000023091) is located on SSC4 and contained a predicted frameshift variant (rs702520311). *PEX19* is assumed to be under regulation by peroxisome proliferator-activated receptor gamma coactivator-1 alpha (PGC-1α) increasing the mitochondrial FA oxidation in human primary myotubes [58]. In addition, peroxisomes are intimately associated with lipid droplets and they are able to perform FA oxidation and lipid synthesis [59]. The frameshift variant was found to be fixed in the Iberian boars for the alternative allele (*PEX19:c.98_102dupAAGTC*), whereas in the Landrace sows the alternative allele was present with a frequency of 0.2.

10. The sterile alpha motif domain containing 4B (*SAMD4B*) gene (ENSSSCG00000016927), located on SSC16, contained a predicted frameshift variant that causes a stop gained (rs709630954). This gene was found to produce leanness and myopathy in mice due to the dysregulation of the rapamycin complex 1 (mTORC1) signalling [60].

Segregation analysis of the selected indels

The ten selected indels were genotyped in 143 BC1_DU, 160 BC1_LD and 138 BC1_PI individuals. Table 4 shows the genotype frequencies of indels in each backcross. Allele genotyping of the *CRP:c.515delT* indel failed and this indel was discarded for posterior analysis.

GWAS results

Nine indels located within genes related with lipid metabolism and genotyped in the three experimental backcrosses were selected for the association analysis. GWAS was performed with a linear-mixed model (*GEMMA* software) among the genotypes of 38,424 SNPs segregating in the three backcrosses and the nine selected indels and the fatty acid composition in muscle.

GWAS results in the merged dataset showed no significant association between the nine genotyped indels and the 20 FA composition traits in IMF. However, a suggestive association between the *C1QTNF12:c.557_559delCCG* indel and the eicosadienoic acid (C20:2(n-6)) (*p*-value = 1.77×10^{-5}, FDR = 5.34×10^{-2}) was identified in the BC1_PI backcross-specific GWAS (Fig 4). This association was not found in the other two backcrosses BC1_DU (*p*-value = 1.65×10^{-1}, FDR = 8.92×10^{-1}) and BC1_LD (*p*-value = 1.63×10^{-1}, FDR = 9.11×10^{-1}) (S2 Fig).

Eicosadienoic acid is the elongated product of linoleic acid, an essential FA that is taken from the diet [61,62] and can be desaturated into arachidonic acid which participates in multiple regulatory pathways [61,62]. The BC1_PI pigs carrying the *C1QTNF12:c.557_559delCCG* allele had a lower proportion of C20:2(n-6). This result was not observed in the rest of backcrosses despite the *C1QTNF12* indel was segregating at similar frequencies in the three backcrosses (Table 4). We hypothesize that other mechanisms could be modulating the levels of C20:2(n-6) in the BC1_DU and BC1_LD backcrosses and masking the effect of the *C1QTNF12* indel.

C1QTNF12 is a gene member of the C1QTNF family which preferentially acts in adipose tissue and liver regulating glucose uptake and fatty acid metabolism [54]. C1QTNF12 can also form heterodimers with the protein encoded by the *ERFE* (erythroferrone) gene, another gene member of the C1QTNF family, which is mainly expressed in skeletal muscle and is able to reduce the circulating levels of free FAs without affecting adipose tissue lipolysis [63].

Table 4. Genotype frequencies of the nine indels found in each backcross. For each backcross, 143 BC1_DU, 160 BC1_LD and 138 BC1_PI were genotyped.

	ASPH			PEX19			C1QTNF12		
Alleles	-/-	ACAG/-	ACAG/ACAG	-/-	AAGTC/-	AAGTC/AAGTC	CCG/CCG	CCG/-	-/-
BC1_DU	0.82	0.17	0.01	0.21	0.48	0.31	0.57	0.43	0
BC1_LD	0.65	0.32	0.04	0.03	0.25	0.73	0.53	0.47	0
BC1_PI	0.92	0.08	0	0	0.03	0.97	0.55	0.45	0

	CCR7			LYST			CAPN9		
Alleles	-/-	A/-	A/A	-/-	CCA/-	CCA/CCA	GAA/GAA	GAA/-	-/-
BC1_DU	0.60	0.35	0.05	0.73	0.22	0.05	0.43	0.57	0
BC1_LD	0.28	0.46	0.26	0.14	0.57	0.29	0.48	0.51	0.01
BC1_PI	0.87	0.13	0	0.28	0.55	0.16	0.54	0.46	0.01

	JMJD1C			GZMA			SAMD4B		
Alleles	CAG/CAG	CAG/-	-/-	-/-	GTT/-	GTT/GTT	TG/TG	TG/-	-/-
BC1_DU	0.54	0.46	0	0.83	0.16	0.01	0.33	0.64	0.04
BC1_LD	0.51	0.49	0	0.17	0.45	0.38	0.05	0.50	0.45
BC1_PI	0.59	0.41	0	0.30	0.64	0.07	0.13	0.48	0.39

Therefore, alterations of the C1QTNF12/ERFE heterodimer may modify the circulation of free FAs and their accumulation in IMF.

Based on the data from the Ensembl project (www.ensembl.org; release 92) using the *Sscrofa11.1* assembly, the porcine *C1QTNF12* gene consists of 8 exons and 7 introns (Ensembl ID: ENSSSCG00000003333). The identified indel produces an inframe deletion of three bases (CCG) in the exon 5 of *C1QTNF12*, which has the consequence of removing the alanine in the position 186 of the final protein. This alanine deletion was located in the C1q/TNF-like domain of C1QTNF12, a domain that is highly conserved among the *C1QTNF12* gene of mammals (Fig 5) and other vertebrate species [64], and is characteristic of the C1QTNF family. Furthermore, the alanine deletion in the position 186 was predicted to cause a new α-helix formation in the secondary structure of C1QTNF12, which could produce an impairment in the protein function (Fig 6).

Nonetheless, the *C1QTNF12* indel was not the most significant genetic variant on SSC6 (Fig 4 and S2 Table). Thus, further studies are required in order to analyse whether other genes or other *C1QTNF12* polymorphisms may be the cause for the differences in the eicosadienoic acid abundance.
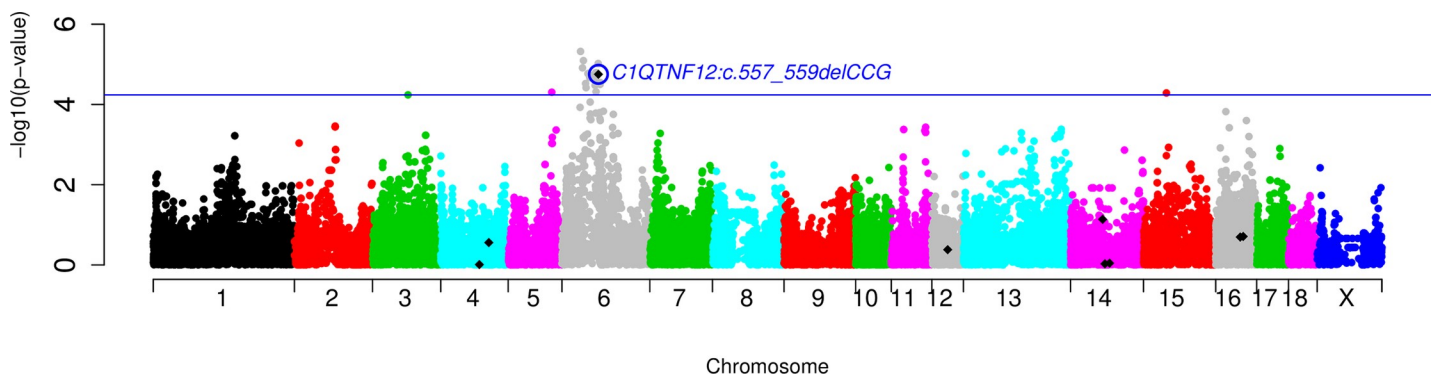

Fig 4. Manhattan plot representing the GWAS analysis for the relative abundance of eicosadienoic acid in the *Longissimus dorsi* muscle of the BC1_PI backcross where the *C1QTNF12* indel (blue circle) was suggestive (FDR≤0.1, blue line). The nine genotyped indels are depicted as black rhombi.
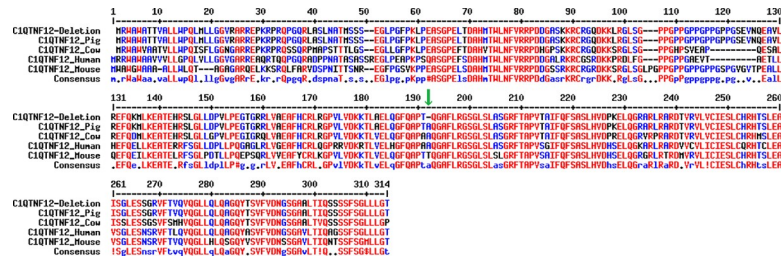
Fig 5. Multiple sequence alignment based on MULTALIN [65] of the porcine C1QTNF12 protein sequence with the deletion and the reference sequences of the C1QTNF12 protein in pig, human, cow and mouse. The green arrow points out the deletion.
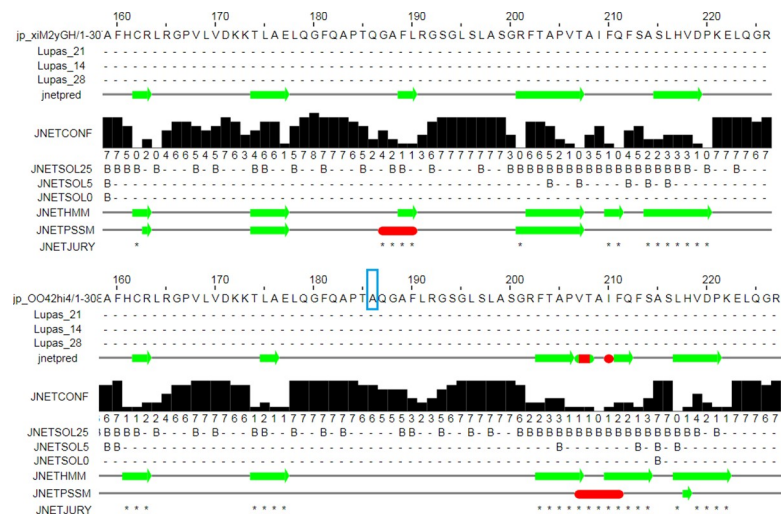
Fig 6. JPred4 prediction of the change in the secondary structure of the porcine C1QTNF12 protein when the alanine in the position 186 (A inside the blue rectangle) of the reference sequence (bottom) is deleted (above). Red segments represent alpha helices and green, beta sheets.

In conclusion, in this study we used three different programs that increased the accuracy of indel detection. Nine indels of the 1.6 million indels detected *in silico* were validated through genotyping in three different backcrosses, showing different allelic frequencies. In addition, a suggestive association was found between the *C1QTNF12:c.557_559delCCG* indel and the eicosadienoic acid abundance. Thus, indels can also be used as genetic markers associated with phenotypic traits of interest.

Supporting information

S1 Fig. Population structure of the three IBMAP experimental backcrosses (BC1_DU, BC1_LD and BC1_PI).
(TIF)

S2 Fig. QQ-plots and Manhattan plots for the relative abundance of eicosadienoic acid in the *Longissimus dorsi* muscle of the merged dataset and the three backcrosses (BC1_DU, BC1_LD and BC1_PI). The nine genotyped indels are depicted as black rhombi and the *C1QTNF12* indel is encircled in blue. Red and blue lines indicate those polymorphisms that

were below the genome-wide significance and suggestive threshold (FDR $\leq$ 0.05 and FDR $\leq$ 0.1, respectively).
(TIF)

S1 Table. Compressed vcf file containing the 1,631,044 indels found in common between the three programs (*Dindel*, *SAMtools mpileup* and *GATK*) used for predicting the consequence type and severity of the indels by the *VEP platform*.
(RAR)

S2 Table. GEMMA output for the suggestive (FDR$\leq$0.1) SNPs found in the GWAS analysis for the log2 normalization of the relative abundance of eicosadienoic acid in the *Longissimus dorsi* muscle of the BC1_PI population.
(CSV)

Author Contributions

Conceptualization: Josep M. Folch, Maria Ballester.

Data curation: Daniel Crespo-Piazuelo.

Formal analysis: Daniel Crespo-Piazuelo, Josep M. Folch, Maria Ballester.

Funding acquisition: Ana I. Fernández, Josep M. Folch.

Investigation: Daniel Crespo-Piazuelo, Lourdes Criado-Mesas, Manuel Revilla, Anna Castelló.

Methodology: Daniel Crespo-Piazuelo, Josep M. Folch, Maria Ballester.

Resources: Ana I. Fernández, Josep M. Folch.

Software: Daniel Crespo-Piazuelo.

Supervision: Josep M. Folch, Maria Ballester.

Validation: Daniel Crespo-Piazuelo, Lourdes Criado-Mesas, Anna Castelló.

Visualization: Daniel Crespo-Piazuelo.

Writing – original draft: Daniel Crespo-Piazuelo.

Writing – review & editing: Josep M. Folch, Maria Ballester.

References

1. Karlsson A, Enfält AC, Essén-Gustavsson B, Lundström K, Rydhmer L, Stern S. Muscle histochemical and biochemical properties in relation to meat quality during selection for increased lean tissue growth rate in pigs. J Anim Sci. 1993; 71: 930–8. https://doi.org/10.2527/1993.714930x PMID: 8478293

2. Estévez M, Morcuende D, Cava López R. Physico-chemical characteristics of M. Longissimus dorsi from three lines of free-range reared Iberian pigs slaughtered at 90 kg live-weight and commercial pigs: a comparative study. Meat Sci. 2003; 64: 499–506. https://doi.org/10.1016/S0309-1740(02)00228-0 PMID: 22063133

3. Serra X, Gil F, Pérez-Enciso M, Oliver M., Vázquez J., Gispert M, et al. A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs. Livest Prod Sci. 1998; 56: 215–223.

4. Wood JD, Enser M, Fisher A V., Nute GR, Sheard PR, Richardson RI, et al. Fat deposition, fatty acid composition and meat quality: A review. Meat Sci. 2008; 78: 343–58. https://doi.org/10.1016/j.meatsci.2007.07.019 PMID: 22062452

5. Poudyal H, Panchal SK, Diwan V, Brown L. Omega-3 fatty acids and metabolic syndrome: effects and emerging mechanisms of action. Prog Lipid Res. 2011; 50: 372–87. https://doi.org/10.1016/j.plipres.2011.06.003 PMID: 21762726

6. Michas G, Micha R, Zampelas A. Dietary fats and cardiovascular disease: putting together the pieces of a complicated puzzle. Atherosclerosis. 2014; 234: 320–8. https://doi.org/10.1016/j.atherosclerosis.2014.03.013 PMID: 24727233

7. Frühbeck G, Méndez-Giménez L, Fernández-Formoso J-A, Fernández S, Rodríguez A. Regulation of adipocyte lipolysis. Nutr Res Rev. 2014; 27: 63–93. https://doi.org/10.1017/S095442241400002X PMID: 24872083

8. Wood JD, Enser M, Fisher A V., Nute GR, Richardson RI, Sheard PR. Manipulating meat quality and composition. Proc Nutr Soc. 1999; 58: 363–70. PMID: 10466178

9. Cameron N. Genetic and phenotypic parameters for carcass traits, meat and eating quality traits in pigs. Livest Prod Sci. 1990; 26: 119–135.

10. Cameron ND, Enser MB. Fatty acid composition of lipid in Longissimus dorsi muscle of Duroc and British Landrace pigs and its relationship with eating quality. Meat Sci. 1991; 29: 295–307. https://doi.org/10.1016/0309-1740(91)90009-F PMID: 22061434

11. Ntawubizi M, Colman E, Janssens S, Raes K, Buys N, De Smet S. Genetic parameters for intramuscular fatty acid composition and metabolism in pigs. J Anim Sci. 2010; 88: 1286–94. https://doi.org/10.2527/jas.2009-2355 PMID: 20042548

12. Casellas J, Noguera JL, Reixach J, Díaz I, Amills M, Quintanilla R. Bayes factor analyses of heritability for serum and muscle lipid traits in Duroc pigs. J Anim Sci. 2010; 88: 2246–54. https://doi.org/10.2527/jas.2009-2205 PMID: 20418459

13. Mullikin JC, Hunt SE, Cole CG, Mortimore BJ, Rice CM, Burton J, et al. An SNP map of human chromosome 22. Nature. 2000; 407: 516–520. https://doi.org/10.1038/35035089 PMID: 11029003

14. Dawson E. A SNP Resource for Human Chromosome 22: Extracting Dense Clusters of SNPs From the Genomic Sequence. Genome Res. 2001; 11: 170–178. https://doi.org/10.1101/gr.156901 PMID: 11156626

15. Weber JL, David D, Heil J, Fan Y, Zhao C, Marth G. Human Diallelic Insertion/Deletion Polymorphisms. Am J Hum Genet. 2002; 71: 854–862. https://doi.org/10.1086/342727 PMID: 12205564

16. Mills RE, Pittard WS, Mullaney JM, Farooq U, Creasy TH, Mahurkar A a., et al. Natural genetic variation caused by small insertions and deletions in the human genome. Genome Res. 2011; 21: 830–839. https://doi.org/10.1101/gr.115907.110 PMID: 21460062

17. Berger J, Suzuki T, Senti K-A, Stubbs J, Schaffner G, Dickson BJ. Genetic mapping with SNP markers in Drosophila. Nat Genet. 2001; 29: 475–81. https://doi.org/10.1038/ng773 PMID: 11726933

18. Wicks SR, Yeh RT, Gish WR, Waterston RH, Plasterk RH. Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. Nat Genet. 2001; 28: 160–164. https://doi.org/10.1038/88878 PMID: 11381264

19. Britten RJ. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. Proc Natl Acad Sci. 2002; 99: 13633–13635. https://doi.org/10.1073/pnas.172510699 PMID: 12368483

20. Britten RJ, Rowen L, Williams J, Cameron RA. Majority of divergence between closely related DNA samples is due to indels. Proc Natl Acad Sci. 2003; 100: 4661–4665. https://doi.org/10.1073/pnas.0330964100 PMID: 12672966

21. Anzai T, Shiina T, Kimura N, Yanagiya K, Kohara S, Shigenari A, et al. Comparative sequencing of human and chimpanzee MHC class I regions unveils insertions/deletions as the major path to genomic divergence. Proc Natl Acad Sci U S A. 2003; 100: 7708–13. https://doi.org/10.1073/pnas.1230533100 PMID: 12799463

22. Ophir R, Graur D. Patterns and rates of indel evolution in processed pseudogenes from humans and murids. Gene. 1997; 205: 191–202. PMID: 9461394

23. Zhang Z, Gerstein M. Patterns of nucleotide substitution, insertion and deletion in the human genome inferred from pseudogenes. Nucleic Acids Res. 2003; 31: 5338–5348. https://doi.org/10.1093/nar/gkg745 PMID: 12954770

24. Fan Y, Wang W, Ma G, Liang L, Shi Q, Tao S. Patterns of insertion and deletion in Mammalian genomes. Curr Genomics. 2007; 8: 370–8. https://doi.org/10.2174/138920207783406479 PMID: 19412437

25. de Jong W, Rydén L. Causes of more frequent deletions than insertions in mutations and protein evolution. Nature. 1981; 290: 157–159. https://doi.org/10.1038/290157a0 PMID: 7207597

26. Molnár J, Nagy T, Stéger V, Tóth G, Marincs F, Barta E. Genome sequencing and analysis of Mangalica, a fatty local pig of Hungary. BMC Genomics. 2014; 15: 761. https://doi.org/10.1186/1471-2164-15-761 PMID: 25193519

27. Chen L, Jin L, Li M, Tian S, Che T, Tang Q, et al. Snapshot of Structural Variations in the Tibetan Wild Boar Genome at Single-Nucleotide Resolution. J Genet Genomics. 2014; 41: 653–657. https://doi.org/10.1016/j.jgg.2014.10.001 PMID: 25527106

28. Kang H, Wang H, Fan Z, Zhao P, Khan A, Yin Z, et al. Resequencing diverse Chinese indigenous breeds to enrich the map of genomic variations in swine. Genomics. Elsevier Inc.; 2015; 106: 286–94.

29. Wang Z, Chen Q, Liao R, Zhang Z, Zhang X, Liu X, et al. Genome-wide genetic variation discovery in Chinese Taihu pig breeds using next generation sequencing. Anim Genet. 2017; 48: 38–47. https://doi.org/10.1111/age.12465 PMID: 27461929

30. Ren Z, Liu W, Zheng R, Zuo B, Xu D, Lei M, et al. A 304 bp insertion/deletion mutation in promoter region induces the increase of porcine IDH3β gene expression. Mol Biol Rep. 2012; 39: 1419–26. https://doi.org/10.1007/s11033-011-0876-1 PMID: 21617947

31. Zang L, Wang Y, Sun B, Zhang X, Yang C, Kang L, et al. Identification of a 13 bp indel polymorphism in the 3'-UTR of DGAT2 gene associated with backfat thickness and lean percentage in pigs. Gene. 2016; 576: 729–33. https://doi.org/10.1016/j.gene.2015.09.047 PMID: 26407871

32. Pérez-Enciso M, Clop A, Noguera JL, Ovilo C, Coll A, Folch JM, et al. A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. J Anim Sci. 2000; 78: 2525–31. https://doi.org/10.2527/2000.78102525x PMID: 11048916

33. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press; 1989. pp. E3–E4.

34. Revilla M, Puig-Oliveras A, Castelló A, Crespo-Piazuelo D, Paludo E, Fernández AI, et al. A global analysis of CNVs in swine using whole genome sequence data and association analysis with fatty acid composition and growth traits. Davoli R, editor. PLoS One. 2017; 12: e0177014. https://doi.org/10.1371/journal.pone.0177014 PMID: 28472114

35. Andrews S. FastQC: a quality control tool for high throughput sequence data. In: Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc. 2010.

36. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168

37. Neuman J a., Isakov O, Shomron N. Analysis of insertion-deletion from deep-sequencing data: software evaluation for optimal detection. Brief Bioinform. 2013; 14: 46–55. https://doi.org/10.1093/bib/bbs013 PMID: 22707752

38. Albers C a., Lunter G, MacArthur DG, McVean G, Ouwehand WH, Durbin R. Dindel: Accurate indel calls from short-read data. Genome Res. 2011; 21: 961–973. https://doi.org/10.1101/gr.112326.110 PMID: 20980555

39. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

40. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20: 1297–1303. https://doi.org/10.1101/gr.107524.110 PMID: 20644199

41. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics. 2010; 26: 2069–2070. https://doi.org/10.1093/bioinformatics/btq330 PMID: 20562413

42. Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4: a protein secondary structure prediction server. Nucleic Acids Res. 2015; 43: W389–W394. https://doi.org/10.1093/nar/gkv332 PMID: 25883141

43. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559–75. https://doi.org/10.1086/519795 PMID: 17701901

44. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012; 44: 821–4. https://doi.org/10.1038/ng.2310 PMID: 22706312

45. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 1995; 57: 289–300.

46. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008; 456: 53–9. https://doi.org/10.1038/nature07517 PMID: 18987734

47. Gealekman O, Gurav K, Chouinard M, Straubhaar J, Thompson M, Malkani S, et al. Control of adipose tissue expandability in response to high fat diet by the insulin-like growth factor-binding protein-4. J Biol Chem. 2014; 289: 18327–38. https://doi.org/10.1074/jbc.M113.545798 PMID: 24778188

48. Walder K, McMillan J, Lapsys N, Kriketos A, Trevaskis J, Civitarese A, et al. Calpain 3 gene expression in skeletal muscle is associated with body fat content and measures of insulin resistance. Int J Obes Relat Metab Disord. 2002; 26: 442–9. PMID: 12075569

49. Cheverud JM, Fawcett GL, Jarvis JP, Norgard EA, Pavlicev M, Pletscher LS, et al. Calpain-10 is a component of the obesity-related quantitative trait locus Adip1. J Lipid Res. 2010; 51: 907–913. https://doi.org/10.1194/jlr.M900128 PMID: 20388922

50. Sano T, Iwashita M, Nagayasu S, Yamashita A, Shinjo T, Hashikata A, et al. Protection from diet-induced obesity and insulin resistance in mice lacking CCL19-CCR7 signaling. Obesity (Silver Spring). 2015; 23: 1460–71.

51. Hellmann J, Sansbury BE, Holden CR, Tang Y, Wong B, Wysoczynski M, et al. CCR7 Maintains Nonresolving Lymph Node and Adipose Inflammation in Obesity. Diabetes. 2016; 65: 2268–81. https://doi.org/10.2337/db15-1689 PMID: 27207557

52. Visser M. Elevated C-Reactive Protein Levels in Overweight and Obese Adults. JAMA. 1999; 282: 2131. PMID: 10591334

53. Enomoto T, Ohashi K, Shibata R, Higuchi A, Maruyama S, Izumiya Y, et al. Adipolin/C1qdc2/CTRP12 protein functions as an adipokine that improves glucose metabolism. J Biol Chem. 2011; 286: 34552–8. https://doi.org/10.1074/jbc.M111.277319 PMID: 21849507

54. Wei Z, Peterson JM, Lei X, Cebotaru L, Wolfgang MJ, Baldeviano GC, et al. C1q/TNF-related Protein-12 (CTRP12), a Novel Adipokine That Improves Insulin Sensitivity and Glycemic Control in Mouse Models of Obesity and Diabetes. J Biol Chem. 2012; 287: 10301–10315. https://doi.org/10.1074/jbc.M111.303651 PMID: 22275362

55. Lindholm-Perry AK, Cunningham HC, Kuehn LA, Vallet JL, Keele JW, Foote AP, et al. Relationships between the genes expressed in the mesenteric adipose tissue of beef cattle and feed intake and gain. Anim Genet. 2017;

56. Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, et al. Forty-Three Loci Associated with Plasma Lipoprotein Size, Concentration, and Cholesterol Content in Genome-Wide Analysis. Abecasis GR, editor. PLoS Genet. 2009; 5: e1000730. https://doi.org/10.1371/journal.pgen.1000730 PMID: 19936222

57. Chico Y, Lafita M, Ramírez-Duque P, Merino F, Ochoa B. Alterations in erythrocyte membrane lipid and fatty acid composition in Chediak-Higashi syndrome. Biochim Biophys Acta. 2000; 1502: 380–90. https://doi.org/10.1016/s0925-4439(00)00062-4 PMID: 11068180

58. Huang T-Y, Zheng D, Houmard JA, Brault JJ, Hickner RC, Cortright RN. Overexpression of PGC-1α Increases Peroxisomal and Mitochondrial Fatty Acid Oxidation in Human Primary Myotubes. Am J Physiol—Endocrinol Metab. 2017; ajpendo.00331.2016.

59. Lodhi IJ, Semenkovich CF. Peroxisomes: a nexus for lipid metabolism and cellular signaling. Cell Metab. Elsevier Inc.; 2014; 19: 380–92. https://doi.org/10.1016/j.cmet.2014.01.002 PMID: 24508507

60. Chen Z, Holland W, Shelton JM, Ali A, Zhan X, Won S, et al. Mutation of mouse Samd4 causes leanness, myopathy, uncoupled mitochondrial respiration, and dysregulated mTORC1 signaling. Proc Natl Acad Sci U S A. 2014; 111: 7367–72. https://doi.org/10.1073/pnas.1406511111 PMID: 24799716

61. Lagarde M, Bernoud-Hubac N, Calzada C, Véricel E, Guichardant M. Lipidomics of essential fatty acids and oxygenated metabolites. Mol Nutr Food Res. 2013; 57: 1347–58. https://doi.org/10.1002/mnfr.201200828 PMID: 23818385

62. Saini RK, Keum Y. Omega-3 and omega-6 polyunsaturated fatty acids: Dietary sources, metabolism, and signi fi cance—A review. Life Sci. 2018; 203: 255–267. https://doi.org/10.1016/j.lfs.2018.04.049 PMID: 29715470

63. Seldin MM, Peterson JM, Byerly MS, Wei Z, Wong GW. Myonectin (CTRP15), a novel myokine that links skeletal muscle to systemic lipid homeostasis. J Biol Chem. 2012; 287: 11968–80. https://doi.org/10.1074/jbc.M111.336834 PMID: 22351773

64. Wei Z, Lei X, Seldin MM, Wong GW. Endopeptidase cleavage generates a functionally distinct isoform of C1q/tumor necrosis factor-related protein-12 (CTRP12) with an altered oligomeric state and signaling specificity. J Biol Chem. 2012; 287: 35804–14. https://doi.org/10.1074/jbc.M112.365965 PMID: 22942287

65. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 1988; 16: 10881–90. https://doi.org/10.1093/nar/16.22.10881 PMID: 2849754